ASCI I/O: SGPFS

Accelerating Scalable Global Parallel File Systems to the Marketplace

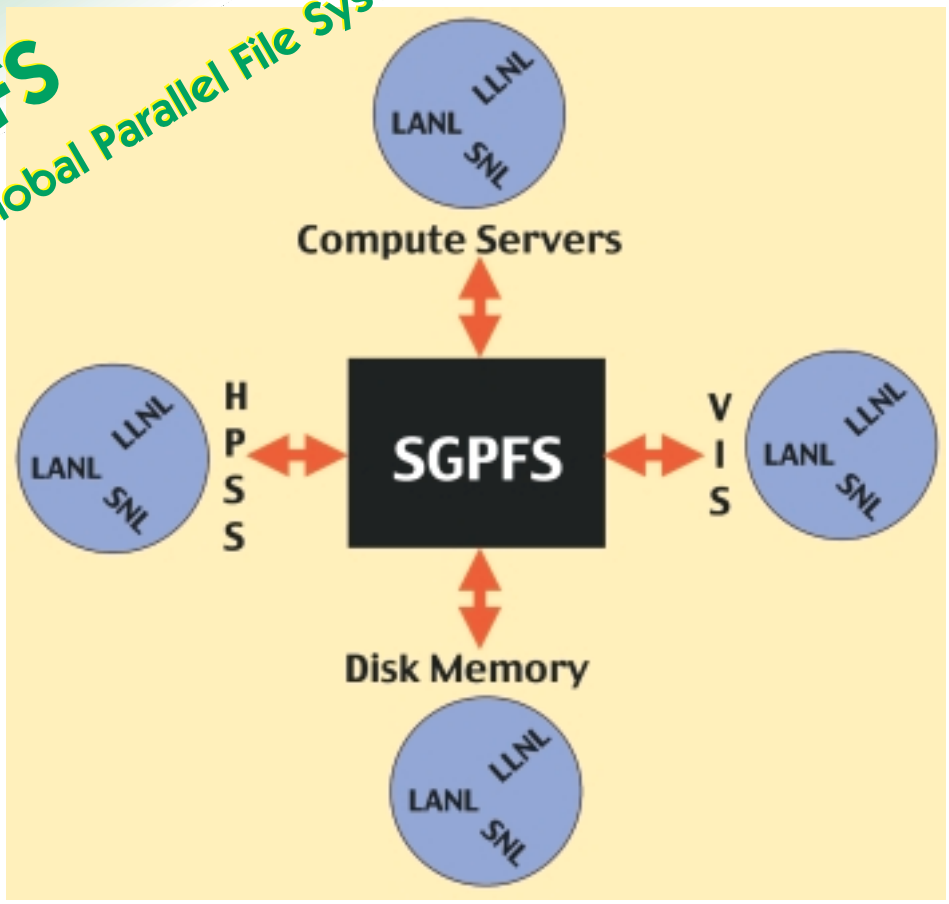

Figure 1. A Scalable Global Parallel File System is seen as a key component to ASCI.

The Scalable Input/Output Project (SIO) has identified a need for an integrated memory hierarchy and I/O solution for the entire ASCI environment from processor registers to archival tape; from mesh generation to visualization. The file system, a key part of the hierarchy, is of special interest to SIO in that it has received little attention in the scale required by ASCI. Project members at the three national defense laboratories involved in the ASCI Program have written a white paper proposing an effort to develop a Scalable Global Parallel File System (SGPFS) with the cooperation of interested parties in industry and academia.

The Need For SGPFS

The imposing computational capabilities needed to support the ASCI Program are especially demanding in the file system area. ASCI sites have a near-term requirement to perform parallel reads and writes to the same file at hundreds of gigabytes per second. Furthermore, it is expedient that this capability be brought to market faster than the current business plans of high-end computer vendors. Therefore, the Department of Energy, in concert with the national laboratories of Los Alamos, Lawrence Livermore, Sandia; and the National Security Agency; are investigating ways to partner with industry and a number of universities to speed up the time to marketplace for an ASCI capable parallel file system. The specific needs for ASCI I/O are shown Figure 2. The hardware to support this I/O strategy is being addressed by industry in a timely manner. The relevant hardware areas include Storage Area Networking and cluster computing technology. The envisioned architecture and infrastructure that will provide a solution are shown in Figure 3. The software is being partly addressed by a number of industry and university efforts including homogeneous parallel file systems, heterogeneous non-parallel global file systems, and MPI-I/O. There is much activity in this area, but it is not converging fast enough for the needs of the ASCI project. The hope for SGPFS is to coalesce the efforts within the industry, university, and government user communities to provide long-term benefits for all parties. To explore the feasibility of such a cooperative effort a SGPFS Industry/User Workshop, was held in Santa Fe, New Mexico on September 23-–24, 1999.

Anticipated Benefits of SGPFS

The benefits to the ASCI and other government user communities are fairly obvious. Global, parallel, and scal-

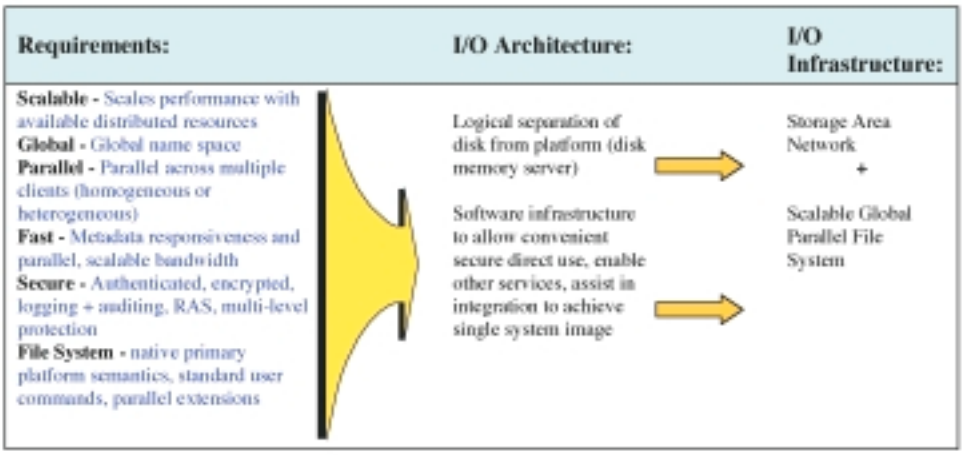

Figure 2. Key functionality for SGPFS.

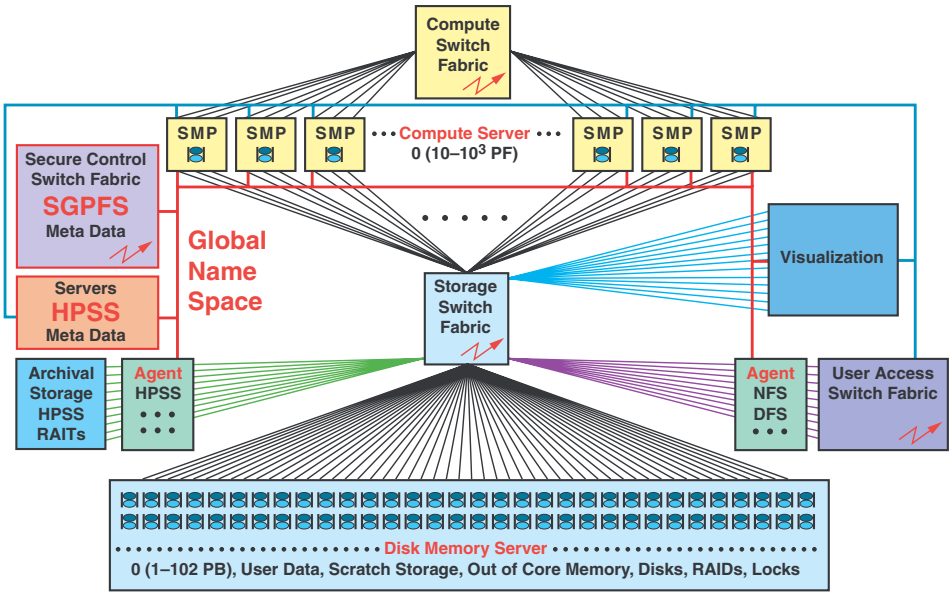

Figure 3. A scalable, resilient architecture is required.

able file systems will be available for current and future use from several vendors on a number of platforms. Necessary security requirements will be met from the beginning. A single system image is enabled for cluster computing. Networks of heterogeneous computing resources will be supported. Several standard interfaces will be supported including Posix , and MPI-I/O.

There are also benefits for many of the industrial partners that are willing to participate in SGPFS. They will be able to satisfy future RFP requirements for high-performance computing. It provides an opportunity to accelerate current or planned activities with some assurance that the effort will not be a dead end. As the general computing environment moves to a clustered computing model in the next several years, SGPFS may provide a standardized solution for an important part of that environment. Since a prototype implementation is expected to be available, vendors will have a way to provide future file system solutions without having to bear the entire cost. The technologies necessary to develop a file system that meets the extensive SGPFS requirements will benefit the entire high-end user community. Facilities to support Reliability, Availability, and Serviceability (RAS) may be applicable to other software systems. Furthermore, a scalable file system will be of interest to both technical computation and commercial computation.

Summary

The ASCI program is investigating ways to accelerate high-performance parallel file systems to the marketplace. These ultra-capable file systems will provide useful technologies for the entire high-end market.

For more information on the ASCI Scalable I/O project, please contact:

Judy Sturtevant, (505)845-9448, jesturt@sandia.gov; or

Tyce McLarty, (505) 667-6034, ttm@lanl.gov; or

Kim Yates, (925) 423-5535, rkyates@llnl.gov

Work performed under the auspices of the U.S. DOE by LLNL under contract No. W-7405-ENG-48